CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

ELSEVIER
CrossMark
Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 35 (2014) 1672 – 1681
Procedia Computer Science

18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014

Study and application of detecting blinks for communication assistant tool

Ippei Torii*, Kaoruko Ohtani, Naohiro Ishii

Aichi Institute of Technology, 1247 Yachigusa, Yakusa-Cho, Toyota, Aichi, 470-0392, Japan

Abstract

We try to develop a support application for physically handicapped children to communicate with others by blinks. Because of limited body movements and mental disorders, many of them cannot communicate with their families or caregivers. If they can use applications in smart phones by blinks, they will be able to tell what they really need or want. Fist, we try to detect an eye area by using OpenCv. Then we develop the way to detect opening and closing of eyes. We combine the method using saturation and using complexity of image to get more accurate results to detect blinks. Then we develop the technique into a communication application that has the accurate and high-precision blink determination system to detect letters and put them into sound.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).
Peer-review under responsibility of KES International.

Keywords: VOCA (Voice Output Communication Aid); Phsically handicapped children; OpenCv; Haar-like eye ditection

1. Introduction

Lately, cell phones and portable game players are gathering attentions as useful communication support tools. Especially so called smart phones that are combinations of cell phones and PDA has been in the limelight. Users can install many applications into smart phones and customize them easily. Those are advantages of smart phones and they have been spread out in the world.

Handicapped children with severe mental disability tend not to have ability of verbal communication, so they need some support tools to express their thoughts or needs. Some assistant applications for PDA (Personal Digital Assistant), such as

* Corresponding author. Tel.: +81-0565-48-8121; fax: +81-0565-48-0277.
E-mail address: mac@aitech.ac.jp

1877-0509 © 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).
Peer-review under responsibility of KES International.
doi:10.1016/j.procs.2014.08.260

Drop Talk[1], Voice4u[2], TapToTalk[3], aimed to help handicapped children who have communication disorder have already introduced. Those communication assistant tools are called VOCA (Voice Output Communication Aid). Although many studies about VOCA had been carried out[4,5,6] and school educational fields have adopted these tools, they are not come into general use because of the high price and complicated operations.

Therefore, we had studied and developed a new VOCA application "Let's Talk!"[7,8,9] In general, a user has to choose the subject, the predicate and the object followed by tapping "talk" on these VOCA applications. In Let's Talk!, a user only tap a button and the sentence will come out with a voice. The simple operations of this application are remarkable and highly supported by users.

We presented our study "Augmentative and Alternative Communication with Digital Assistant for Autistic Children" at the international conference of ESPA 2012[10]. This application was highly approved by many participants because of the low price and the easy manipulation. Some of the researchers suggested this application would be useful for people with mental disorders.

Special support schools needed some support tools for physically handicapped children. We will study and develop a support tool for physically handicapped children based on the system we have developed in Let's Talk!

2. Purpose of the study

On this paper, physically handicapped children are defined as children with permanent disablements with a trunk and limbs because of cerebral palsy, muscular dystrophy, spina bifida and so on.

Their body movements are very limited and most of them need to stay in beds all the time. Many of them also have mental disorders, so they cannot communicate with their families or caregivers. It prevents helpers to understand what they really need or think.

The most common way to communicate with those physically handicapped children is using "Yes=○" or "No=×" cards. (Figure 1) For example, if a caregiver wants to ask a child whether he/she feels hot or cold, the caregiver will ask him/her "Are you hot?" and show him/her boards with ○ and × by turns. If the child will take a look at ○ board or put his/her eye on it longer than the other one, the caregiver will know he/she may feel hot.

Fig. 1. Using "Yes=○" or "No=×" cards to communicate with a physically handicapped child

But caregivers have to predict what patients need or want to say by their experiences or circumstances in this method. So the questions made by caregivers can be totally different from what patients really want to say. And also, it is difficult to figure out the movement of eyes of patients. Sometimes caregivers have to just guess the answer.

Some communication support tools using movements of eyes for these physically handicapped people have been released already, such as TalkEye[11] or Let's Chat[12] in (Figure 2). For example, TalkEye requires the executive head set to measure the movement of eyes. It is difficult to introduce those tools to public schools because of the high prices.

Fig. 2. "TalkEye" and "Let's Chat"

Based on the above points, we try to develop an application for physically handicapped children to tell their needs or wants with blinks.

Students in special support schools in Japan could feel the sense of accomplishment that was reported on their own through by using our former communication support application Let's Talk! They learned how to communicate with others and became to have motivation to relate with others through this application. We also want physically handicapped children to feel the sense of accomplishment and the joy of communication. We use iPad (By Apple Inc. iOS5.0) from Apple for this study.

Generally, the cost will be very high if you ask a company to develop an application. This will raise the price of an application. We think it is very important for academic institutions to return results of our studies to society. This is one of the ways for us to contribute to the education field to study and develop useful applications free of charge. So we aim to develop a new application for physically handicapped children and provide it at no charge.

We use a built-in camera of a smart phone as an input devise for easy and fast processing of picture data. It will help to rise up the accuracy of processing data. We fix a smart phone on a bed to be able to use it with lying down position (Figure 3).

A built-in camera of a smart phone is lower sensitive comparing with a high-performance CCD camera. So if we try to process pictures continuously in real time, it will freeze because of the heavy data. Especially it is difficult to detect eye areas in considerably dark places, such as schools or hospitals. Therefore, to rise up the accuracy of processing of picture data and to detect blinks without freezing will be two most important aims for this study.

After we finish to develop this application, we will correct clinical data, bugs and needs from users by using the network we have established with special support schools all over Japan. Then we will improve this application based on the data.

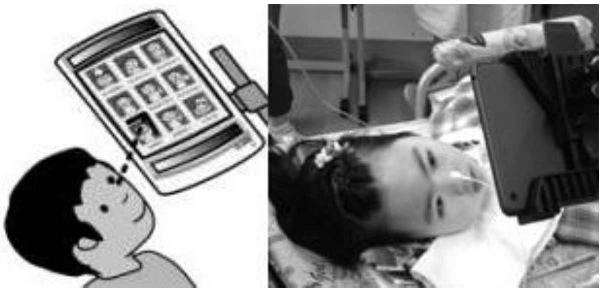

Fig. 3. Using the application with lying down position

3. Structure of the system

Process of detecting blinks is performed in the following steps,

① To detect an eye area (By using Opencv Haar-like eye-detection)

② To distinguish opening and closing of eyes (By using the complexity of binarized image)

③ To add the method using saturation to detect blink (Aiming more accurate detection)

④ To decide by a conscious blink (To define what is a "conscious blink")

⑤ To improve the accuracy of detection of a blink

3.1. Detection of an eye area

There are many methods to detect an object. We choose OpenCv that is a library of programming functions for real time computer vision for image processing in this study. OpenCV's face detector uses a method that Viola, P. [13] developed first and Lienhart, R. [14] improved.

Figure 4 shows the image of Haar-like eye-detection. An eye area is darker than a cheek area when the picture of face is shaded off (Figure 4①). Therefore, if we cut off a rectangular area that is peripheral of eyes and calculate averages of brightness of the upper part and the lower part separately, the upper part is darker than the lower part (Figure 4②). And brightness of the top of a nose is brighter than sides of a nose (Figure 4③). We use these differences of brightness to figure out a face area and to remove other areas. After about 20 times of the removing process, we use the last one as a face data.

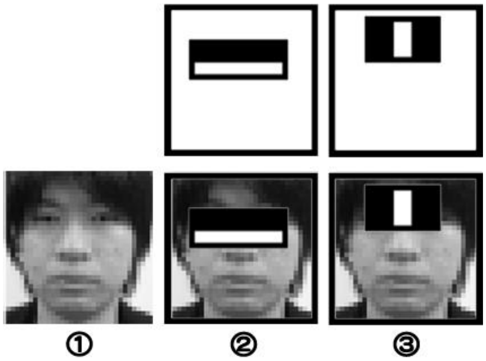

Fig. 4. Image of Haar-like features

Next, we detect an eye-area. Since an eye-area is inside a face area, we use data of a face area obtained by the method to obtain the average brightness of upper and lower part of eyes to define an eye-area.

3.2. Detection of eye opening and closing

There are two methods to detect a blink of eyes. First one is calculating the size of the black area of eyes by using spiral labeling[15] and considering it as a blink when the size becomes smaller than the threshold. The second one is using the difference of the value of brightness of images to determine a blink. Figure 5 shows the image of spiral labeling. In spiral labeling, we search pixels from the starting pixel (1 in a square in Figure 5), calculate the medial level of the difference, and count the pixels within the threshold. In this method, we can reduce the time of processing if we can get the starting pixel because the area of processing is limited.

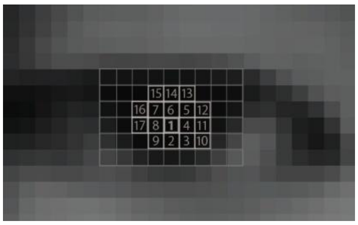

Fig. 5. Image of spiral labeling

But many physically handicapped children tend not to be able to open their eyes wide enough and their iris of the eye is relatively small, so it is difficult to detect the center of iris of the eye. So we use Value in HSV (Hue, Saturation, Value) in color space[16] to determine a blink. In this method, we cut an eye area based on the coordinate data obtained by Haar-like classifier and size it to reduce the load to a devise.

Then we obtain the average of brightness of the eye area. When someone closes his/her eyes, the average will rise. We determine it as eye closing. Figure 6 shows the image of eye opening and closing.

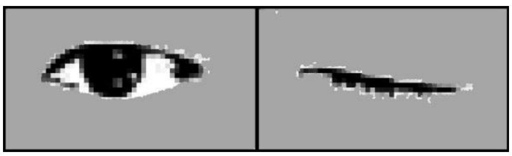

Fig. 6. Image of eye opening and closing

Usually, the classrooms of the most of special support schools are relatively dark to avoid giving extra impetuses to children, so we cannot get enough amount of light. It is difficult to determine a blink, because there is no difference of darkness around the eye area and the average of brightness of white and black part of eyes is close to the brightness of skin.

3.3. Detecting method by saturation in color space

Here we develop the method using saturation in color space (Torii et al. 2013). In this method, we calculate the average of saturation (0 to 255 in saturation of HSV) of area C (the center of the iris) and W (white part of the eye) in Figure 7. If the measured value is lower than the average of saturation, we determine it as eye closing. We collect many numbers of eye area data to calculate the average of saturation for eye opening .

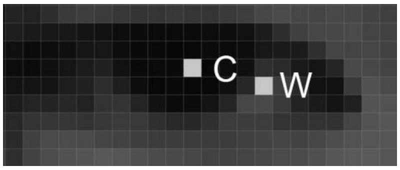

Fig. 7. To Obtain the average of saturation

When a user starts this application, it will calibrate the picture of eye opening. Based on this picture, the application determines it as eye closing when the saturation of eye area is lower than the threshold. When a camera on a tablet starts, auto focus and brightness adjustment will be automatically used. As shown in Figure 8, if auto focus and brightness adjustment are automatically performed, the difference of the calibration data arises every time the application starts.

Fig. 8. Difference of the data between auto focus and brightness adjustment

So we stop to calibrate the eye area when the application starts. We use the threshold (the average of saturation) between 5 flames to 14 flames before of the present flame to determine eye closing. We change the average of saturation (19.50) into the average of HSV (0 to 255). And we try to find the least failure value based on this average S=49. The method using the saturation has high sensitivity. But on the other hand, it will detect a little movement such as an unconscious blink or moving of the face. So we add complexity of image of the eye in last 3 flames including the present flame to exclude the error to determine a blink.

3.4. Detecting method by complexity of image

In this method, we use the difference of the outlines between eye opening and closing[17]. We determine it as eye closing when the pixel of the difference of amount of edge in the picture becomes flat. The threshold is based on the average value of the amount of edges of 10 flames (5 flames to 14 flames before of the present flame). We determine a blink with the difference between the threshold and the present flame. Figure9 shows the utilization to each frame of the process complexity and saturation of blink detection.

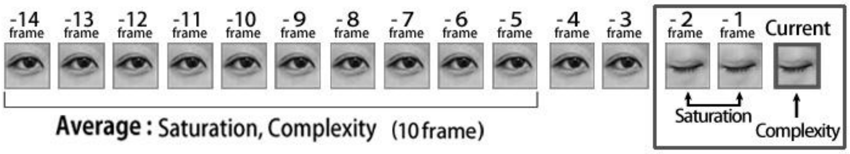

Fig. 9. Complexity of the image of the eye

The purpose of adding this method is to stop to detect an unconscious blink or a movement of the face. Figure 10 shows the image of complexity of the image of the eye.

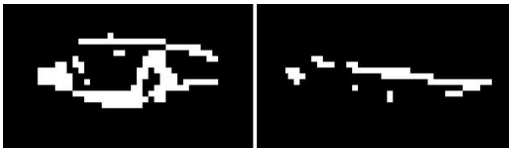

Fig. 10. Complexity of image of the eye

First, we try to find the most optimum value of combination of the saturation and complexity to determine eye closing. When we set the saturation (BP) at 49 (S=49), an error will occur when the value is bigger than 89% and it will not detect the conscious blink when the value is lower than 85%. When we set the number of flames (Time) more than 4, the application does not count an unconscious blink but the error to detect blinks increases. When the complexity (LP) is lower than 73%, it does not detect a blink and make an error when LP is higher than 78%. The setting that has less error is Case4 (LP=0.75, DP=0.86). We set this number as the threshold.

Figure 11 shows the image when the application starts. ABC shows the correlation values of complexity and DEF shows the correlation value of saturation. A is the value of complexity of the present flame. B is the average value of complexity of 5 to 14 flames before of the present flame. C is the value of complexity multiplying the threshold LP (0.75) to the value of the present flame. D is the value of saturation of the present flame. E is the average value of saturation of 5 to 14 flames before of the present flame. F is the value of saturation multiplying the threshold DP (0.86) to the value of the present flame. When B is larger than A, and E is larger than D, we determine it as eye closing.

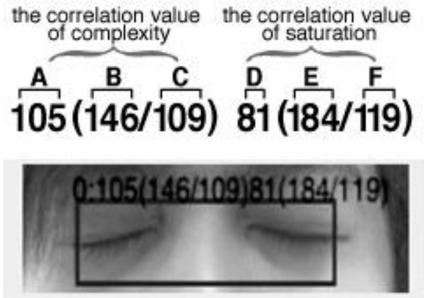

Fig. 11. Image and value of eye area

3.5. Developing afterimage method

After we developed the basic model of the application, we carried on a clinical experiment with eyes closed determination using the complexity in the school for handicapped children in Japan and we could get certain results. However, subjects suffering from spinal muscular atrophy have very weak muscle strength and can not blink longer enough. So we need to improve the application to allow us determine blinks even situations of users are different.

For users who can blink strongly enough, the method to determine blinks by the complexity and saturation is appropriate, but another detection method is required for users who can not blink strongly enough like this subject suffering from spinal muscular atrophy. So we need to detect the unconscious blinks that we excluded. Therefore, it is necessary to increase the processing power for detection of blinks.

Since the tablet is fixed to the bed, we detect the position of eyes using OpenCV only when we start the application. It reduces the processing load and we can use 4 to 25 frames per second to determine blinks.

We use afterimages to determine weak and fast blinks instead of complexity and saturation. When the area (C, Figure12) that is overlapping part of the afterimage of iris (A, Figure12) and the current iris situation (B, Figure12) is diminishing rapidly comparing from a few frames before, we determine it as closing of eyes. It is possible to exclude a slow change and capture a rapid change.

By the method of using the afterimage, the number of the past frames to compare increases and we can capture changes with high accuracy. As a result, we can quantify the changes of series of movements "eye-opening → eye-closing → eye-opening". And it is possible to determine the situation of eyes as short blink, long blink, closing eyes continuously, except the malfunction due to fine movement of eyes. Figure12 shows determination by afterimages visually. A is the afterimage to be compared, B is the iris in current frame, and C is the overlapped area of the afterimage and current frame.

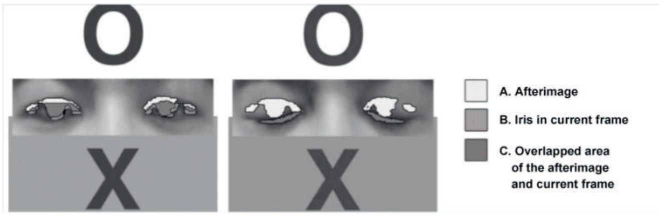

Fig. 12. Detection of unconscious blink by afterimages, visually.

In order to handle with face color and brightness of light, calibration is automatically activated when the black part in the image is doubled. Calibration means to detect the position of eyes and to set a threshold for binarizing. The value of binarization threshold is too high or too low, it can not detect the blinking. We perform calibration to obtain the correct value for blink determination and set the threshold. As a result, it is possible to automatically capture the position of eyes in both a dark place and a bright place or even if a user moves his/her face.

The upper graph in Figure13 is a change of the difference of values between the number of pixels of the current frame and afterimage. The middle line is the threshold and we set to determine it as a blink when more than 1/18 of the total number of pixels is changed. When the value drops below the threshold after the value exceeds the threshold then certain number of frames (setting to five frames) are passed, it is determined as a blink. The vertical straight lines indicate the determination of blinks. It is possible to set the parameters not to react with too long or too fast blinks.

The bottom part in Figure13 is a graph for the calibration. This shows the change of fractions of the pixel volume in the current frame. The dotted line is the threshold when binarizing (Binarization Threshold). The numbers are up and down automatically by the brightness of the room and the color of the skin to create an optimal state for detection. The middle line is the ideal value of calibration. When the state of opening eyes is close to the line, it is working properly.

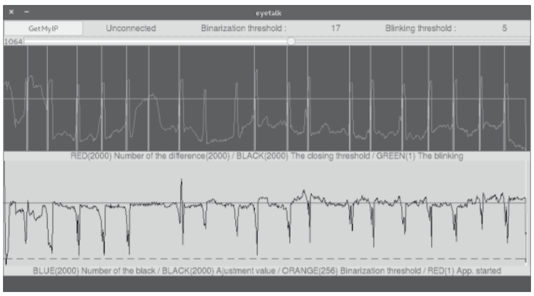

Fig. 13. The graph of change of the difference of values between the number of pixels of the current flame and afterimage

It is possible to detect blinks using the value of the current frame, but if a user has a long eyelashes, the outline becomes complicated because it acquires the contour of the eyelashes, the value is unstable when the eyes are closing. Therefore, we designed the method of using the afterimage to correspond to blinks of any user.

Fig. 14. New applications using character table of Japanese for blink determination to detect letters and put them into sound

We developed the application into a communication application that has the accurate and high-precision blink determination system to detect letters and put them into sound. We named it "Eye Talk!" We try to make the application to be able to select a character in the least number of times. Figure15 shows the image of the application. A user chooses a consonant in "column" first and a vowel in "row" next from the character table of Japanese of this application.

Fig. 15. Items of Command

After choosing a letter, the frame cursor is moving along "Command" items, which includes Voiced/Semi-voicedsound symbols, Delete a letter, Select, Sound and Delete All, on the top of the screen one by one. We considered the visibility and usability to decide the order of Command items and the moving speed of the frame cursor. If a user blinks when the frame cursor is on Select, the letter is chosen. By continuing this operation, a user can create a sentence as long as he/she needs. Then the voice sound comes out when a user choose Sound button. We are adding new functions such as the prediction conversion, the input history and the mail function that a user can send a text by e-mail.

A user needs to practice to abstain the operations of Eye Talk and it is quite difficult for persons with intellectual disabilities. So we developed a modified communication application to indicate needs simply by "Yes = ○" and "No = ×" by blinks. We named this application "Eye Tell"(Fig. 16). A user can make original symbols suitable for the level of handicap. A lot of parents and teacher use Eye Tell and special support education schools in Japan approve of the concept of the application.

First, a helper makes 2 symbols that are suitable for the situation of a user in this application. Then the symbols on right and left sides of the screen turn on and off reciprocally. A user chooses the symbol when it turns on by a blink. The application judges the blink and put the sentence into voice sound. A helper can make symbols freely by picture, sound and text. Twenty one sets of symbols can be stored. If a user acquire how to operate the application by blinks, he/she will be able to use "Eye Talk", that is required more complicated operation.

Fig. 16. Eye Tell

4. Conclusion

Our study is not just a methodology. The main purpose of our study is to apply the technology of visualization of information to a daily life. Especially we aim to develop a useful application, to research and correct the data, to modify the problems and to apply the methods to the next step.

We are on the process of filing a patent for the high quality blink detection method and distributing it free of charge in Japan. For further development of this method, we research and develop the method of determination of gazing directions by analyzing the eye movement. If we can detect the gazing direction accurately on the tablet, burden on a user is reduced. It is applicable to any application by limiting the processing range to achieve higher speed and simplifying the process and high detection accuracy.

To determine gazing directions, we use the afterimage method we have developed in this study. To determine gazing directions, we measure the rate of increase or decrease of the area of the white part of eyes by the afterimage method for processing an image of the eyes. In this way, we can determine gazing directions with high accuracy and less error. By combined the method to determine gazing directions with the blink detection method, a user will be able to get more unfettered communication.

The support applications are living necessities for handicapped people. They should be practical and easy to use. If we know what physically handicapped children want to tell us with this application, we can understand each other more. It will encourage the physically handicapped people to communicate with others and will lead the society to support each other regardless of handicap.

References

1. DropTalk, *http://droplet.ddo.jp/droptalk/*
2. Voice4u, *http://voice4uaac.com/jp/*
3. TapToTalk, *http://www.taptotalk.com/index.aspx*
4. K. Nakamura. Effects and limitations of communication aids. Journal of communication aids for people with cognitive disabilities. *Higher brain function research*; 2001.p.194-200.
5. S. Sakai. VOCA application to Autistic Children. *National Welfare Foundation for Disabled Children*; 1999.p. 29-32.
6. K. Nakamura. Use of VOCA in Children with Intellectual Disabilities and/or Autistic Tendencies : Observations of VOCA Play Settings and Speech Abilities in Two Special Schools. *The Japanese*.
7. I. Torii, K. Ohtani, T. Niwa and N. Ishii. Voice Output Communication Aid Application for Personal Digital Assistant for Autistic Children. *IIAI/ACIS International Symposium on Innovative E-Services and Information Systems (IEIS 2012)*; 2012. p. 329-333.

8. I. Torii, K. Ohtani, T. Niwa and N. Ishii. Development and Study of Support Applications for Autistic Children. *ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (IEEE ACIS SNPD 2013;2013* p. 420-425.
9. I. Torii, K. Ohtani, T. Niwa and N. Ishii. Study and Development of Schedule Application for Autistic Children. *IIAI International Conference on Advanced Applied Informatics (IIAI AII ICSCAI 2013)*;2013. p. 251-255.
10. I. Torii, K. Ohtani, T. Niwa, A.Yamamoto and N. Ishii. Augmentative and Alternative Communication with Digital Assistant for Autistic Children. *IEEE International Conference on Emerging Signal Processing Applications（IEEE/ ESPA Las Vegas 2012）*;2012. p. 71~74.
11. TalkEye, *http://www.takei-si.co.jp/product/talkeye.html*
12. Let's Chat, *http://panasonic.co.jp/phc/products/home/communicationaids.html*
13. Paul Viola and Michael J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. *IEEE CVPR*; 2001.
14. Rainer Lienhart and Jochen Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection. *IEEE ICIP*; 2002, Vol. 1, p. 900-903.
15. A. Tsukada. Automatic Detection of Eye Blinks using Spiral Labeling. *Symposium on Sensing via Image Information(SSI 2003)*; 2003. vol.9, p. 501-506.
16. G. H. Joblore and D. Greenberg. Color Spaces for Computer Graphics. *Computer Graphics*. vol.12. p.20-27.
17. I. Torii, K. Ohtani, T. Niwa and N. Ishii. Study and Development of Support Tool with Blinks for Physically Handicapped Children. *IEEE 25th International Conference on Tools with Artificial Intelligence (ICTAI 2013)*; 2013. p. 116-122.